

QoS based Admission Control using Multipath Scheduler for IP over Satellite Networks

Lukman Audah¹, Zhili Sun², Haitham Cruickshank³

¹Wireless and Radio Science Centre (WARAS), Faculty of Electrical Engineering,
Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, Malaysia

^{2,3}Institute for Communication Systems (ICS), University of Surrey, GU2 7XH, Surrey, United Kingdom

Article Info

Article history:

Received Jun 28, 2017

Revised Sep 12, 2017

Accepted Sep 30, 2017

Keyword:

Admission control

IP over satellites

Multipath scheduler

Network simulator (NS-2)

Quality of service (QoS)

ABSTRACT

This paper presents a novel scheduling algorithm to support quality of service (QoS) for multiservice applications over integrated satellite and terrestrial networks using admission control system with multipath selection capabilities. The algorithm exploits the multipath routing paradigm over LEO and GEO satellites constellation in order to achieve optimum end-to-end QoS of the client-server Internet architecture for HTTP web service, file transfer, video streaming and VoIP applications. The proposed multipath scheduler over the satellite networks advocates load balancing technique based on optimum time-bandwidth in order to accommodate the burst of application traffics. The method tries to balance the bandwidth load and queue length on each link over satellite in order to fulfil the optimum QoS level for each traffic type. Each connection of a traffic type will be routed over a link with the least bandwidth load and queue length at current time in order to avoid congestion state. The multipath routing scheduling decision is based on per connection granularity so that packet reordering at the receiver side could be avoided. The performance evaluation of IP over satellites has been carried out using multiple connections, different file sizes and bit-error-rate (BER) variations to measure the packet delay, loss ratio and throughput.

*Copyright © 2017 Institute of Advanced Engineering and Science.
All rights reserved.*

Corresponding Author:

Lukman Audah,
Wireless and Radio Science Centre (WARAS),
Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia,
Parit Raja, 86400 Batu Pahat, Johor, Malaysia.
Email: hanif@uthm.edu.my

1. INTRODUCTION

Satellite communication has evolved tremendously since the launched of Ka-Sat to the space orbit in 2010. The satellite services provide ubiquitous Internet connectivity to most people around the world with broadband digital multimedia contents. Internet users are provided with the high-speed triple-play services with guaranteed QoS.

Although the satellite network has successfully overcome the digital divide in communication system, it still has limitation in term of bandwidth constraint. This limitation will become the main obstacle to support the future Internet services (e.g. 5G and beyond) which consist of billions of connected heterogeneous devices. This issue has been discussed since a long time ago and many solutions have been implemented. One of the suggested method is through efficient bandwidth sharing across the satellites network system. The efficient bandwidth sharing concept adopts the load balancing technique to redistribute the traffic load (e.g. link bandwidth utilization) across the entire network through efficient multipath routing. Previous related studies on load balancing technique [1] only focusing on the Low Earth Orbit (LEO) satellite network using the low delay routing algorithm. Other previous related studies [2-4] applied load balancing

using efficient multipath routing across multiple orbital layers of LEO and Geosynchronous Earth Orbit (GEO) satellite networks. The study has efficiently reduce the traffic congestion within the satellite networks through routing techniques but the burden of routes selection and packet processing are placed on the satellite nodes. Furthermore, the studies did not take into consideration the specific QoS requirements of each traffic type for the routing selection.

Another previous related study on multipath routing over satellites by Taleb et al [5] has proposed a novel method called “Explicit Load Balancing” (ELB) which is a cooperative routing strategy for LEO satellites constellation that enables neighboring satellites to explicitly exchange information on their current congestion status for better distribution of packet traffics over the entire satellite constellation. The method also optimally prevents both congestion and packet drops at the satellites.

Inspired by the idea of ELB method developed by Taleb et al, we proposed an alternative approach of multipath routing over satellites which can support any type of orbital satellites constellation (e.g. LEO, MEO and GEO). We still hold the idea that traffic classification and traffic conditioning could be done in the terrestrial network in order to relieve the satellite workload and prevent scalability problem. In addition, we suggest that the current congestion status over the satellite networks could be shared among the Ground to Satellite Link terminals (GSL) and the routing scheduling decision over which satellite networks should be chosen for traffic diversion will be managed by one or more admission control node. The admission control node maintains two separate list of current traffic load over the satellite networks and the queue length of ground to satellite links. The load balancing of packet traffic from ground to satellite links will be done based on those lists in order to maintain optimum QoS of ‘delay-sensitive’ and ‘throughput-sensitive’ Internet application packet streams. Our scope of research only sees from end-to-end satellite-terrestrial network system perspective. We omit the routing type used among the satellites within the same constellation network and assume that each satellite network maintains their own routing tables. The next sections explain in details of our proposed QoS method for multipath routing over satellite-terrestrial networks and the system performance evaluation using NS-2 network simulations.

2. PROPOSED RESEARCH METHOD

The proposed novel QoS method supports multiservice applications over the 10Mbps of high speed satellite broadband using Ka-Sat like satellite and the Big-LEO systems. The method adopts an admission control system which consists of a Diffserv queue interface and a multipath scheduler in order to regulate, classify and select the optimum multipath routes for the Internet Protocol (IP) packet traffics. In addition we proposed a simple priority queue with selective packets drop function as the satellite OBP. The main focus of this study is developing the multipath routes selection algorithm for the admission control system in order to achieve optimum QoS performance for the multiservice applications traffics. The applications traffics used in the studies are HTTP web, large file transfer using FTP, Video streaming and bidirectional VoIP. The proposed QoS framework has been tested against multiple new connection rates, average server-response files sizes and bit-error-rate (BER) variation using NS-2 simulation software.

The NS-2 network simulation scenario is shown in Figure 1 which consists of 4 remote servers, 4 remote clients, an admission control system, 8 ground to satellite links terminals (GSL), a Geosynchronous satellite (GEO) and 66 Low Earth Orbit (LEO) satellites constellation. The remote servers are connected to bottleneck admission control node using bidirectional link of 20Mbps with 8ms of link delay at the server side while a bidirectional link of 2.5Mbps with 2ms link delay is used to connect each GSL terminal with the client node. The next subsections explain the details about network elements parameters.

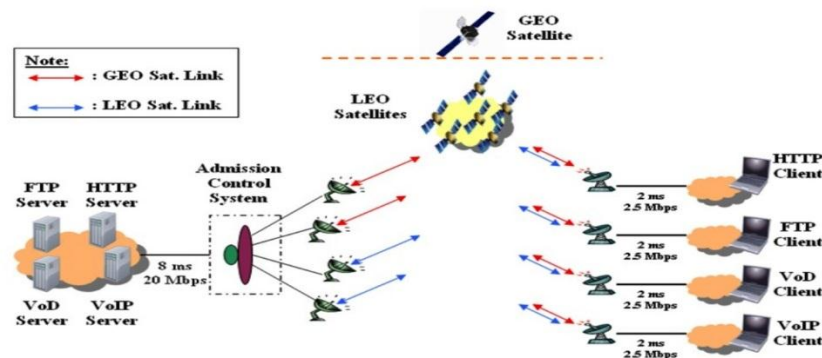


Figure 1. Simulation scenario in NS-2

2.1. Admission Control System

The admission control system will route the traffic connection to multiple paths over both LEO and GEO satellites networks using load balancing method. The method will fully utilized both LEO and GEO satellites networks resources in order to achieve optimum QoS for the throughput-sensitive (e.g. HTTP web, FTP) and delay-sensitive (e.g. VoIP, video streaming) traffics. The load balancing method is based on the following parameters.

- Current load bandwidth of each branching link.
- Current queue length of each branching link.
- Current end-to-end delay based on data collected by the moving agent.

Based on Figure 1, the admission control system consists of the following subcomponents:

- Multipath scheduler (the purple color oval object).
- Diffserv queue interface (the green color circle object).

In NS-2, the multipath scheduler is a node element derived from the Class Node. We extend the basic node functionality to support multipath route selection at run-time by adding the multipath classifier object as shown in Figure 2. The node is designed in such a way that it would only suitable to function as a branching node immediately before the GSL terminals. This makes it capable to regulate the traffic flows over the satellite networks.

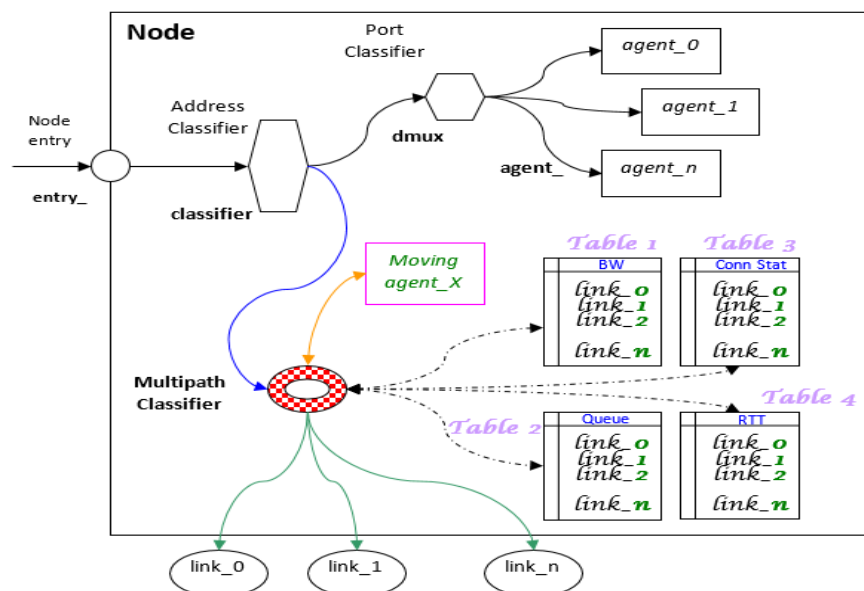


Figure 2. A novel multipath scheduler node structure in NS-2

The multipath scheduler tasks involve measurement of the load bandwidth and queuing buffer size of the forward branching links based on the IP packet forwarding history and keeps its own statistical tables. The tables contain lists of continuously updated current traffic load over the satellite networks and the queue length of ground to satellite links.

In addition, it maintains a routing table based on the unique connection ID for every TCP/UDP connection that passed through it. Every newly established connection unique ID will be registered in the routing table list along with the chosen optimum forwarding route. Subsequent packets from the same connection will follow the same route until the connection is terminated. When the connection is terminated, it will be deregistered from the routing table list to leave space for the subsequent new connection. The registration and deregistration processes provide information to the multipath scheduler about the total live connections and connections termination rates in each branching links over the satellite networks.

Moreover, it probes all the possible paths towards destination for the link delay by periodically creating and dissipating moving agents that continuously moving forward to destination network element and backward to itself. The advantages of using moving agent in the multipath routing over LEO satellites network have been highlighted by Rao et al [6]. We adopted the method in our study as the additional function to the multipath scheduler element for optimum traffic route selection. The moving agent is defined

as small chunks of packets in NS-2 created to record the round-trip-time (RTT) time-stamp for each branching links from the multipath scheduler towards client nodes. The timestamp records both forward and reverse delays between server and client sides. Congestion could be detected if the recorded forward delay increased from previous recorded values or the probe packets dropped as the results of timeout. The forward delay data will be used for comparison with other recorded list of parameters stored in the tables in order to determine the optimum traffic route towards client node.

The optimum route selection for each traffic type is based on the measured data and Diffserv queue information. Packets from the same connection will not be separated in order to avoid packets reordering at the receiver side. The multipath routes forwarding algorithm decides on how the multipath scheduler will handle the current connection flows and the newly established connection flows for each traffic type based on the time-bandwidth sharing concept. The time-bandwidth sharing concept defines the concept of links sharing among lower delay links (i.e. links using LEO satellites network) and higher delay links (i.e. links using GEO satellite network).

2.2. Satellite Networks Configuration

The satellite networks used in the proposed network scenario are based on the Ka-Sat like satellite system located at coordinate 9° east and LEO-Iridium constellation. There are 4 remote servers that transmit multiple TCP and UDP connections to 4 remote clients via 8GSLs. We also introduced a random error model to simulate the satellite network transmission loss characteristics. The error model produced 2 different BER values which are 10^{-7} and 10^{-6} for 2 different error scenarios.

Based on Figure 1, the 4GSLs at the servers side are located at the same latitude and longitude coordinate $(37.9^{\circ}, -122.3^{\circ})$ while the other 4GSLs at the clients side are located at $(42.3^{\circ}, 71.1^{\circ})$. The GEO satellite is located at the coordinate $(0^{\circ}, -100^{\circ})$. The “-” sign refers to location in West region in earth. The LEO satellites constellation are scattered around the globe at low earth orbit. The RTT between the GEO satellite and the respective GSL at server and client side could be calculated as follows [7]:

The relative coordinate difference between GEO satellite and a GSL terminal at the server side is shown as follows:

$$\begin{aligned} \text{GEO} - \text{GSL (server)} &= (0, -100) - (37.9, -122.3) \\ &= (-37.9, 22.3) \end{aligned}$$

The relative coordinate difference between a GSL terminal at the client side and the GEO satellite is shown as follows:

$$\begin{aligned} \text{GSL (Client)} - \text{GEO} &= (42.3, 71.1) - (0, -100) \\ &= (42.3, 171.1) \end{aligned}$$

Then the R_E , h and r parameters are specified as follows:

$$\begin{array}{lll} R_E & - \text{Earth radius from earth's core to earth terminal} & = 6378.137\text{Km} \\ h & - \text{The satellite altitude from the earth equator} & = 35786.6\text{Km} \\ r & - \text{Radius from the earth's core to GEO satellite} & = R_E + h = 42164.737\text{Km} \end{array}$$

Distance between the GEO satellite and GSL terminal can be calculated using the following Equation [7]:

$$R^2 = R_E^2 + r^2 - 2R_E r \cos \theta \quad (1)$$

where

$$\cos \theta = \cos L \cos \varphi \cos l + \sin \varphi \sin l \quad (2)$$

The L and l parameters are the latitude and longitude of the earth coordinate system respectively while the φ parameter is the azimuth angle. As for GEO satellite, we have $\varphi = 0$, thus $\cos \varphi = 1$ and $\sin \varphi = 0$.

The value of $\cos \varphi$ between GEO – GSL (server) is calculated as follows:

$$\cos \varphi_1 = \cos(-37.9^{\circ}) \cos(22.3^{\circ}) = 0.73$$

$$\varphi_1 = 43.11^{\circ}$$

The value of $\cos \varphi$ between GSL (Client) – GEO is calculated as follows:

$$\begin{aligned}\cos \varphi_2 &= \cos(42.3^{\circ})\cos(171.1^{\circ}) = -0.7307 \\ \varphi_2 &= 136.95^{\circ}\end{aligned}$$

Therefore the distance between GEO satellite and GSL terminal (server) can be computed as follows using (1):

$$\begin{aligned}R_1^2 &= (6378.137)^2 + (42164.737)^2 - 2(6378.137)(42164.737)(0.73) \\ &= 1.4259 \times 10^9 \text{ Km}^2 \\ R_1 &= 37761.148 \text{ Km}\end{aligned}$$

Besides that, the distance between GSL (Client) terminal and GEO satellite can also be computed as follows:

$$\begin{aligned}R_2^2 &= (6378.137)^2 + (42164.737)^2 - 2(6378.137)(42164.737)(-0.7307) \\ &= 2.2116 \times 10^9 \text{ Km}^2 \\ R_2 &= 47027.2643 \text{ Km}\end{aligned}$$

Thus the one-way propagation delay from GSL (server) – GEO satellite – GSL (client) can be calculated as follows:

$$\begin{aligned}T_p &= \frac{R_1 + R_2}{c} \\ &= \frac{(37761.148 + 47027.2643) \times 1000}{3 \times 10^8} \\ &= 283 \text{ ms}\end{aligned}$$

Therefore, the *RTT* value over GEO satellite is obtained as follows:

$$RTT = 2T_p = 2 \times 283 \text{ ms} = 566 \text{ ms}$$

Moreover, the *RTT* over LEO network could not be calculated as previously mentioned because we did not monitor closely the coordinate of LEO satellites involved directly in the data transmission. However, an approximation value could be considered as previously calculated in [7] which is $\approx 75.162 \text{ ms}$ with $RTT \approx 150.324 \text{ ms}$.

2.3. Multiservice Application Traffic Modeling

There are 4 applications services that will be used in the proposed network scenario as in Figure 1 which are the generic HTTP web traffic, large files transfer using FTP, Video streaming using Constant Bit Rate (CBR) traffic and bidirectional VoIP using GSM.AMR codec. The end-to-end QoS performance will be measured against variable connection rates and BER values during the entire 15 minutes of simulation time.

2.3.1. HTTP Web Traffic

The HTTP web traffic application in the NS-2 multipath simulations used Pareto distribution to model the server response file size with average value of 300Kbytes. The HTTP web traffic used TCP New Reno as the underlying transport protocol. The TCP segment size used is 1500bytes (i.e. 1460bytes payload + 40bytes header) with maximum congestion window size of 29 packets. The congestion window is set to 29 packets because HTTP packet may travel on both LEO and GEO satellites networks at any possible time. Since the *RTT* over LEO is much smaller than the *RTT* over GEO (e.g. approximately 142 ms tested for the network scenario without BER), the maximum TCP congestion window size value is obtained based on the

maximum TCP throughput Equation (W_{max}/RTT) using the 2.5Mbps of the satellite uplink/downlink bandwidth as the benchmark and 1500bytes of TCP segment size.

2.3.2. Large File Transfer using FTP

The FTP traffic application used in this study is for the Internet large file transfer with average file size value (e.g. $avg_{-}(x)$) of 5Mbytes based on Pareto distribution. The FTP traffic used TCP New Reno as the underlying transport protocol. The TCP segment size used is 1500bytes (i.e. 1460bytes payload + 40bytes header) with maximum congestion window size of 43 packets.

The congestion window is set to 43 packets because FTP packet may travel over GEO satellites networks at most of the simulation time based on the defined routing scheduling algorithm. Since the RTT over GEO is higher than the RTT in LEO (e.g. more than 500ms tested for the network scenario without BER), the maximum TCP congestion window size could reach more than 100 packets when calculated using 2.5Mbps of the satellite uplink/downlink bandwidth as the benchmark, 500ms RTT and 1500bytes of TCP segment size. However, 100 packets of congestion window size are equal to 150Kbytes and more than the standard maximum TCP window size which is 65535bytes [8]. Too large TCP window setting may introduce error-recovery problem and degrades the performance substantially [9]. Therefore, the maximum TCP congestion window size used in the NS-2 simulations is limited to 43 packets only.

2.3.3. Video Streaming Traffic

The VoD video streaming traffic is modelled using Constant Bit Rate (CBR) traffic generator in NS-2. We follow the standard MPEG-transport stream (MPEG-TS) packet size specification used in the Integrated Services Digital Broadcasting - Terrestrial (ISDB-T), Digital Video Broadcast (DVB-T/C/S) and Advanced Television Systems Committee (ATSC) which is 208bytes (188bytes transport stream packet + 20bytes of FEC data) [10]. The 188bytes packet size is a standard payload size chosen for compatibility with ATM system while the additional 20bytes is used for the Reed-Solomon forward error correction.

2.3.4. Voice over IP Traffic

The bidirectional VoIP application traffic used in the multipath simulations used GSM.AMR codec. The average new VoIP connection rate between two conversation entities varies between 1 and 30 connection/minute based on the Exponential distribution. Moreover, the average conversation duration per connection is 10 minute/connection based Pareto distribution. Similarly with other traffics, the new connection is created regardless of the completion of previous connection.

2.4. Differentiated Services (Diffserv) Configuration

The Diffserv queuing system in multipath simulations used Random Early Detection (RED) queue type and the Time Sliding Window 3 Color Marker (TSW3CM) of policer type. The total buffer size of a physical queue is 500 packets with average packets size of 1500bytes. The 3 virtual queues are virtually some fractions of the physical queue size which corresponds to the minimum threshold ($minTh$) and maximum threshold ($maxTh$). Assuming that 90% of the total buffer size used for user traffics, therefore the $maxTh$ could be set equally to all traffic type which is 450 packets.

The CIR and PIR values are set based on 10Mbps of the total branching links bandwidth (e.g. 2.5Mbps of the satellite uplink/downlink). This is done in order to limit the traffic rate that pass through the satellite networks so that the total rate does not exceed far beyond 10Mbps.

2.5. Satellite On-Board Processing (OBP)

The OBP system used in the NS-2 simulations is exactly the same as mentioned in [11] which managed the traffic flows based on delay-sensitive and throughput-sensitive categories. Each satellite in LEO and GEO networks has the OBP queuing system with maximum buffer size of 500 packets with average size of 1500bytes for each packet.

2.6. Delay Scheduling and Queue Function

The term delay scheduling generally means to delay the specific scheduling tasks at some period of time in order to preserve optimum flow of the queuing tasks system. The study of delay scheduling topic has been done by many researchers in the fields of TCP/IP protocol layers for efficient user data locality management and for minimizing the effect of transmission delays over heterogeneous networks [12-14]. Inspired by the work in the previous studies, we adopt the method as part of the admission control system in order to regulate the traffics flow across the terrestrial-satellite networks.

The admission control system at the branching node element also adopts the delayed scheduling algorithm function at its queuing buffer in order to reduce packets dropped of TCP traffic. The method will

delay the TCP packets at random period of time when the queue buffer exceeds some specific capacity thresholds. The delay is done by using a virtual queue on top of the physical queue at the admission control node as shown in Figure 3. When the physical queue towards the branching links have reached the threshold value, any TCP packet that arrives after that will be queued in the virtual queue buffer until the physical queue length drops below the threshold value. Once the physical queue length drops below the threshold value, the TCP packets will pass through over the satellite networks. This is done repeatedly at the admission control node whenever the physical queue length exceeds the threshold. The delayed scheduling may only take split seconds because the admission control will terminate any new arrival connection when the queue threshold is exceeded until the load bandwidth drops below 90% of the link capacity. The delay scheduling and connection termination methods create double protection on the ongoing TCP live connections from packets dropped.

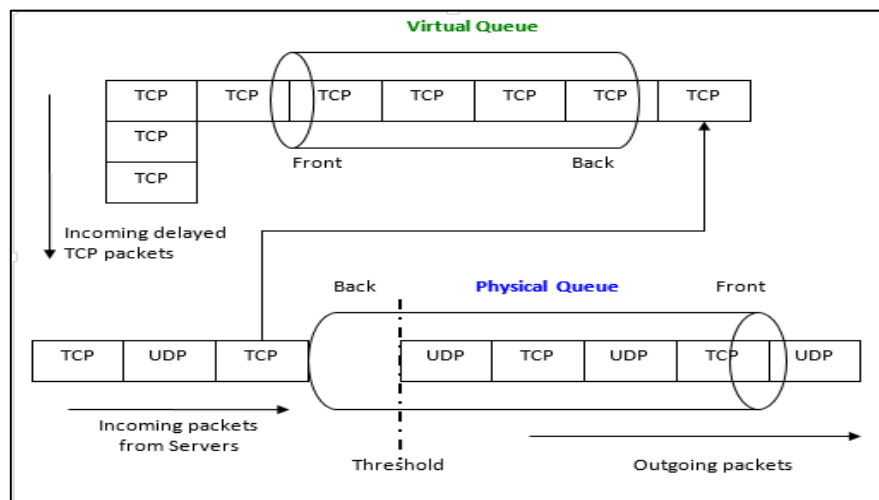


Figure 3. Delay scheduling at admission control node

3. RESULTS AND ANALYSIS

The simulations study comprised of 4 traffic types which are the HTTP web, Large File transfer using FTP, Video streaming and VoIP. In addition, the branching links bandwidth is set to 2.5Mbps each and the satellite links network are exposed to the bit-error-rate. Simulations are done in order to investigate the QoS parameters variations of the proposed system as the results of the increment of new connection rate and also the BER values.

As mentioned in previous section, there are 4 application traffics involved which are HTTP web, large file transfer using FTP, Video streaming and VoIP. All of the traffics start at the same time at 10 second and stop at the end of simulation time which is 900 second. The simulations are done 30 times in order to investigate the effect of new connections arrival rates between 1 and 30 connection per minute. Each point in the following graphs represents an average values calculated during the entire 900 second of simulation time. Although the Pareto and Exponential model distributions are used in are used in all traffic types for certain purposes like inter-arrival time, file size distribution and live-time connection distribution, each traffic type used different set of random values provided by the distribution model because each traffic type used different random number distribution seed.

3.1. Average End-to-End Delay

The packet delay is measured by subtracting the packet received time at the client (t_r) to the packet sending time from server (t_s). The average delay (D) is measured by summing up all packets delays and then divided by the total number of successfully received packet (P_t) at the client side as shown in (3).

$$D(s) = \frac{\sum_{i=1}^{i=N} (t_r - t_s)_i}{P_t} \quad (3)$$

The average end-to-end packet delay in second as shown in Figure 4 is proportional to the increment of average connection rates from each traffic type. The delays are higher at most of the time in BER 10^{-6} scenario and also during higher connection rates mainly because of many packets dropped and network congestion. The global queuing delay and also the invocation of delay scheduling upon the throughput-sensitive traffics are the main factors toward higher delay during network congestion. Besides that, the average delay perceives by the throughput-sensitive traffics are higher at most of the time because the admission control system preferred the routing over GEO network when the routes over LEO network are congested. In addition, the higher delays are also due to the TCP reliable transmissions characteristics like window resizing and packets retransmission. The FTP traffic perceive highest average end-to-end delay values compared to the HTTP web traffic mainly because the traffic carries large file sizes over the network and the which takes longer time to complete transmission and error recovery per connection during congestion periods. The VoIP traffic perceive the lowest delay during the entire simulations times because the proposed admission control node limits the queuing threshold and drops the packets when the queue exceeds threshold limit. In addition, the underlying UDP transport protocol did not do any packet retransmission during network congestion which causes the end-to-end delay lower than TCP traffics. The maximum end-to-end delay for the throughput-sensitive traffic could be seen in BER 10^{-6} simulation scenario which is 3064.1942ms for the FTP traffic while the maximum end-to-end delay for the delay-sensitive traffic is 707.1763ms perceived by the Video streaming traffic. Besides that, the maximum end-to-end delay for HTTP web and VoIP traffic are 933.1949ms and 85.195ms respectively. The proposed system has successfully managed to control the delay according to priority among traffic types and also to maintain the delay of VoIP traffic under the maximum standard recommendation.

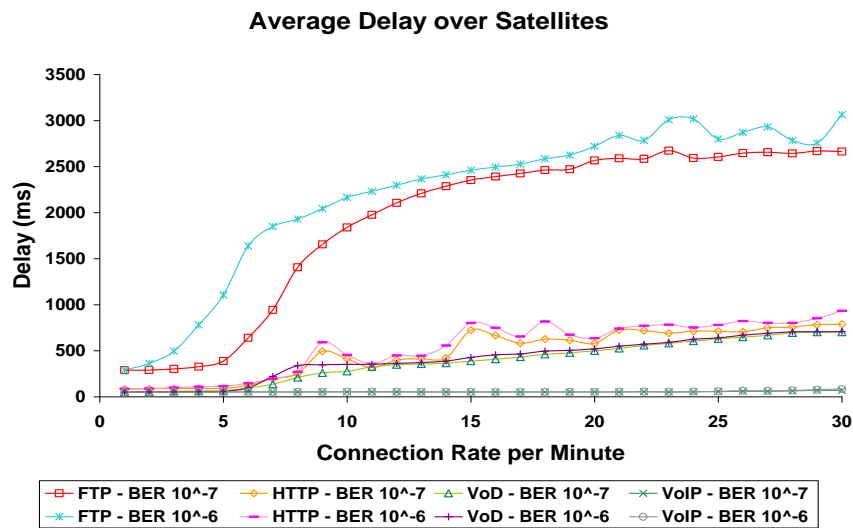


Figure 4. Average end-to-end delay over satellite

3.2. Average End-to-End Loss Ratio

Average packet loss ratio (L) refers to the ratio of total packet loss (P_l) over total transmitted packet from server to client (P_s). Equation (4) shows the loss ratio calculation.

$$L = \frac{\sum_{i=1}^{i=N} (P_l)_i}{\sum_{i=1}^{i=N} (P_s)_i} \quad (4)$$

Based on Figure 5, the end-to-end loss ratio is proportional to the increment of connection rate. The loss ratio is significant in both graphs after 20th connection rates. The loss ratio over satellites scenario is higher mainly due to the link-loss environment. The loss ratio is higher in BER 10^{-6} compared to the one in BER 10^{-7} due to many packets dropped in the satellite networks. The Video streaming traffic has the highest

packet loss in both graphs mainly because of the huge packets streams across the network elements. The traffic streams constant bit-rate traffic for an average of 5 minutes and this may quickly congested the global network system. In addition, the higher loss ratio over satellite networks is also due to the packets early dropped by the admission control when the queue buffer exceeds some threshold limits. The Video streaming traffic is set to have lower priority compared to the VoIP which makes it more preferable to be dropped by the proposed admission control system. The VoIP traffic has the second highest loss ratio over the satellite networks followed by the FTP and HTTP. Although the proposed controlled system successfully produces the optimum end-to-end delay, the trade-off would be higher loss ratio to the delay-sensitive traffics. Moreover, Figure 5 proves that the proposed admission control system has successfully lowered the loss ratio of throughput-sensitive traffics compared to the delay-sensitive traffics for optimum end-to-end QoS. The maximum loss ratio among delay-sensitive traffics over satellites is coming from the Video streaming traffic which corresponds to the value of 0.5247. In addition, the maximum loss ratio among throughput-sensitive traffics is from the FTP traffic with the value of 0.2442 over satellites with BER 10⁻⁶. Moreover, the maximum loss ratio for HTTP web traffic is 0.0139 while the loss ratio for VoIP traffic is 0.3476 in the satellite with BER 10⁻⁶.

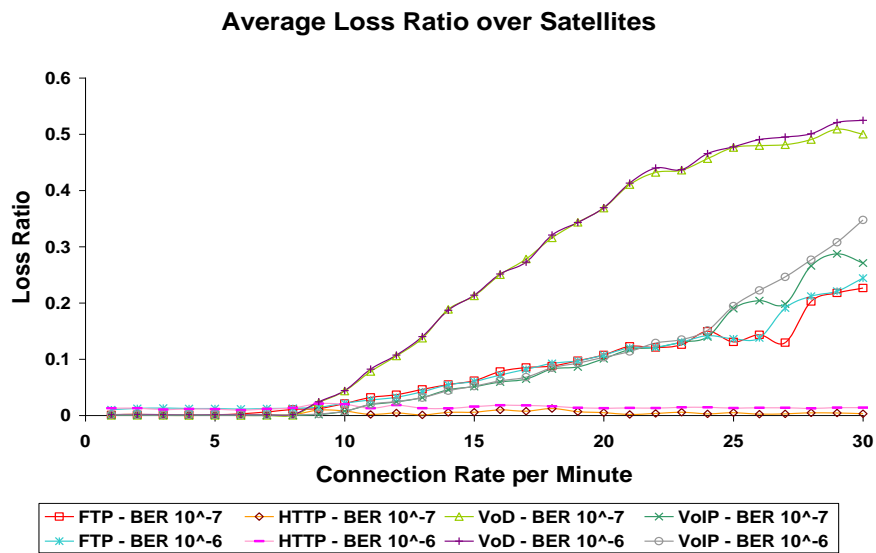


Figure 5. Average end-to-end loss ratio over satellites

3.3. Average End-to-End Throughput

Flow throughput is calculated by dividing the total received packet bytes (P_b) over the duration of a FTP flow connection. The FTP flow duration calculated by subtracting the receiving time of last packet at the client (t_l) to the sending time of first packet of a flow at the server side (t_f). Then, the average flow throughput (B) in bps is calculated by summing up all completed flow throughputs and divided by the total number of completed flows (f_t) as in (5).

$$B = \frac{\sum_{i=1}^{i=f_t} \left(\frac{P_b \times 8}{t_l - t_f} \right)_i}{f_t} \tag{5}$$

Figure 6 shows the average connection throughput for HTTP, web, large file transfer using FTP, Video streaming and VoIP against new connection rates per minute and BER variations. The end-to-end flow throughputs on Figure 6 are closely related to the previous mentioned delay and loss ratio simulation results. The graph also shows that throughputs are inverse proportional to the increment of average new connection rates and also the BER variations over satellites. In general, the throughputs over satellite networks are low. The lower throughputs could be related to the higher end-to-end loss ratios over the satellites due to the link-loss environment and early packets dropped by the proposed admission control system. In addition, higher

global delay in the proposed network scenario due to the queuing delay, packet retransmission and delayed scheduling also contribute to the lower throughputs.

The end-to-end throughputs for the Video streaming and VoIP traffics are lower than the HTTP web and FTP traffics at all time mainly because the traffics are designed with lower maximum generated throughputs. As an example, the Video streaming traffic could reach a maximum throughput of approximately 47.5Kbps while the maximum throughput for VoIP traffic is approximately 5.5Kbps. The minimum throughput for Video streaming traffic over satellites is approximately 22Kbps during worst case scenario. In addition, the minimum throughput for VoIP is 3.4Kbps over the satellite networks in worst case scenario. Unlike the delay-sensitive traffics, the throughput-sensitive traffics are severely affected to the new connection rate and BER variations due to the TCP reliable transmission characteristics. The retransmission process and also the global queuing delays during network congestion have greatly reduced the end-to-end throughputs. The throughput for HTTP web traffic over satellite networks varies between approximately 540Kbps at best case scenario and 212Kbps at worst case scenario. Moreover, the throughputs for FTP traffic vary between approximately 688Kbps and 12.5Kbps over the satellite networks. The throughputs of FTP traffic are lower than the HTTP web traffic at most of the time mainly due to the large file transfer over the network system. Since the random inter-arrival of new connection does not wait until the previous connection has completed, the large file transfer using TCP may quickly congest the network at higher new connection rates and BER due to packets retransmission process in order to successfully complete the data transfer. In addition, both of the throughput-sensitive traffics are imposed with the delayed scheduling over the satellite networks scenario which makes the throughputs further less.

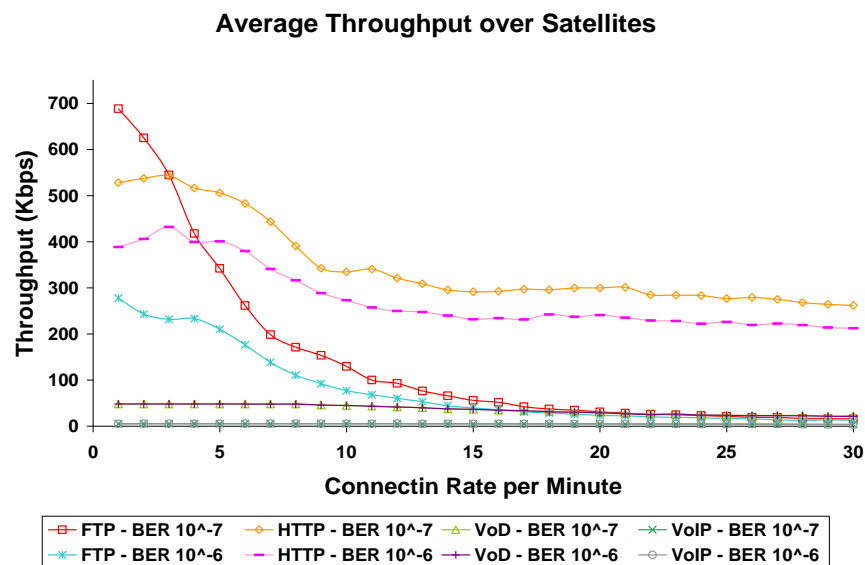


Figure 6. Average end-to-end throughput over satellites

4. CONCLUSION

In this paper, we have done the network layer enhancement using a novel load balancing algorithm method with multipath forwarding over LEO and GEO networks in order to optimize the bandwidth utilization. An admission control with Diffserv queuing system has been placed in the terrestrial network to regulate and differentiate traffic flows in real-time based on the current load bandwidth and queue length in order to reduce the satellite workload for data processing. In addition, we have implemented the delay scheduling function at the admission control node in order to reduce the TCP packets dropped during network congestion. Furthermore, the satellite's on-board-processing system (OBP) is implemented as a simple priority queue for further processing by prioritizing the traffic flow within the satellite network. The proposed system aims to achieve optimum end-to-end QoS for both throughput-sensitive and delay-sensitive traffic types. Moreover, we further improved the admission control system with the delay-scheduling mechanism which main task is to delay the throughput-sensitive traffic at certain period of time during congestion period in order to prevent packet losses before being transmitted over the satellite networks. Finally, the proposed network system has tested in big-scale simulations to produce the QoS parameters that

vary according to the connection rate and BER variations. The simulations produce the end-to-end QoS parameters like delay, loss ratio, throughput and connection statistics which have been plotted into graphs for comparison.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Education Malaysia under Research Acculturation Grant Scheme (RAGS) and Universiti Tun Hussein Onn Malaysia for the generous financial support.

REFERENCES

- [1] X. Liu, X. Yan, Z. Jiang, C. Li and Y. Yang, "A low-complexity routing algorithm based on load balancing for LEO satellite networks," in *Proceedings of IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Boston, USA, September 2015.
- [2] Y. Wang, M. Sheng, K.S. Lui, X. Wang, R. Liu, Y. Zhang and D. Zhou, "Tailored load-aware routing for load balance in multilayered satellite networks," in *Proceedings of IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Boston, USA, September 2015.
- [3] L. Jiang, G. Cui, S. Liu, W. Wang, D. Liu and Y. Chen, "Cooperative relay assisted load balancing scheme based on stackelberg game for hybrid GEO-LEO satellite network," in *Proceedings of IEEE International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1-5, Nanjing, China, October 2015.
- [4] H. Nishiyama, D. Kudoh, N. Kato and N. Kadowaki, "Load balancing and QoS provisioning based on congestion prediction for GEO/LEO hybrid satellite networks," in *Proceedings of the IEEE*, vol. 99, no.11, pp. 1998-2007, November 2011.
- [5] T. Taleb, D. Mashimo, A. Jamalipour, N. Kato and Y. Nemoto, "Explicit load balancing technique for N GEO satellite IP networks with on-board processing capabilities," *IEEE/ACM Transaction on Networking*, vol. 17, no. 1, pp. 281-293, April 2009.
- [6] Y. Rao and R.C. Wang, "Agent-based load balancing routing for LEO satellite networks," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 54, no. 17, pp. 3187-3195, December 2010.
- [7] Z. Sun, *Satellite networking: Principles and protocols*, 1st ed., (John Wiley & Sons Ltd, West Sussex, England, 2005), pp. 32-34.
- [8] V. Jacobson, "TCP extensions for high performance," IETF Request for Comment (RFC) 1323, May 1992.
- [9] M. Allman, S. Floyd and C. Partridge, "Increasing TCP's initial window," IETF Request for Comment (RFC) 3390, October 2002.
- [10] "Information technology – Generic coding of moving pictures and associated audio information: Systems," International Standard, ISO/IEC 13818-1:2000, December 2000.
- [11] L. Audah, Z. Sun and H. Cruickshank, "QoS evaluation of multiservice applications over integrated satellite-terrestrial networks," in *Proceedings of 3rd International Congress of Ultra Modern Telecommunications and Control Systems, ICUMT 2011*, Budapest, Hungary, October 2011.
- [12] M. Markovitch and G. Scalosub, "Bounded delay scheduling with packet dependencies," in *Proceedings of IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)*, pp. 257-262, Toronto, Canada, July 2014.
- [13] Z. Xu and S. Guan, "A low-time-delay scheduling algorithm for video on demand service over P2P networks," in *Proceedings of International Conference on Information and Communications Technologies (ICT 2014)*, Nanjing, China, May 2014.
- [14] A. Abbasi and M. Ghaderi, "Minimum delay scheduling with multi-packet transmission in wireless networks," in *Proceedings of IEEE 16th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1-9, Boston. USA, June 2015.

BIOGRAPHIES OF AUTHORS



Lukman Audah was born in Kuala Lumpur, Malaysia. He received Bachelor of Engineering (Telecommunications) from Universiti Teknologi Malaysia in 2005. He then received his MSc. in Communication Networks and Software from University of Surrey, United Kingdom. He also received PhD in Electronic Engineering from University of Surrey. He is currently a lecturer in the Communication Engineering Department, Universiti Tun Hussein Onn Malaysia. His research interests are wireless and mobile communications, Internet traffic engineering, network system management, data security and satellite communications.



Zhili Sun is a Professor (Chair of Communication Networking) with the Institute for Communication Systems (ICS), University of Surrey, UK. He received his PhD in Computer Science from Lancaster University, UK in 1991. He has been principal investigator and technical coordinator in many research projects funded by European Framework Programme, UK Electronics and Physical Sciences Research Council (EPSRC), European Space Agency (ESA) and industries. He has published over 150 papers in International journals including many papers in the IEEE Transactions, magazines, Journals and conferences and 3 books as author or co-author. He has also served as TPC committee members in many international conferences including the IEEE ICC and Globecom, and been ERSRC College Member, Reviewer and Panel member. His research interests include IP networking protocols and technologies, satellite communications and networking, Internet and teletraffic engineering, network security, mobile and wireless communications and mobile operating system



Haitham Cruickshank works in Institute for Communication Systems (ICS) since January 1996 on several European research projects in the ACTS, ESPRIT, Ten-Telecom and IST programmes. His main research interests are network security, satellite network architectures, VoIP and IP conferencing over satellites. He is currently working in several FP6 projects such as SATLIFE, EuroNGI, and SATNEX. He also teach in the Data and Internet Networking and satellite communication courses at University of Surrey. He is a chartered engineer and corporate member of the IEE in UK. He is also a member of the Satellite and Space Communications Committee of the IEEE ComSoc. I am active in the ETSI BSM (Broadband Satellite Multimedia) and the IETF MSEC groups. In addition, he is the Vice Chair of the COST 272 activity, which is part of the European COST research programme.